DATE: 05/17/2001

TIME: 15:34:08

```
Input Set : A:\Pto.amc
                     Output Set: N:\CRF3\05172001\1830160.raw
        <110> APPLICANT: YLIHONKO, Kristiina
              TORKKELL, Sirke
      6
              PALMU, Kaisa
              HAKALA, Juha
       <120> TITLE OF INVENTION: GENE CLUSTER INVOLVED IN NOGALAMYCIN BIOSYNTHESIS,
              AND ITS USE IN PRODUCTION OF HYBRID ANTIBIOTICS
     10
     12 <130> FILE REFERENCE: 1574/49849
C--> 14 <140> CURRENT APPLICATION NUMBER: US/09/830,160
C--> 14 <141> CURRENT FILING DATE: 2001-04-23
     14 <150> PRIOR APPLICATION NUMBER: PCT/FI99/00870
     15 <151> PRIOR FILING DATE: 1999-10-20
     17 <160> NUMBER OF SEQ ID NOS: 18
     19 <170> SOFTWARE: PatentIn version 3.0
     21 <210> SEO ID NO: 1
     22 <211> LENGTH: 16020
     23 <2:12> TYPE: DNA
     24 <213> ORGANISM: Streptomyces nogalater ATCC 27451
     26 <220> FEATURE:
     27 <221> NAME/KEY: misc_feature
     28 <222> LOCATION: 3799..3800
     29 <223> OTHER INFORMATION: "overlapping sequence in the genes snoaM and snogN"
    .31 <221> NAME/KEY: misc_feature
     32 <222> LOCATION: 6334..6356
     33 <223> OTHER INFORMATION: "overlapping sequence in the genes snoaG and snogC"
     35 <221> NAME/KEY: misc_feature
     36 <222> LOCATION: 13201..13300
     37 <223> OTHER INFORMATION: "unknown region"
     39 <400> SEQUENCE: 1
                                                                                60
     41 agatetegte egecagtgee teggtgaeeg geaacgagee ettggegtag eegagatggg
                                                                               120
     43 agáaaccygt catggtgtgc acgggccagg gataactgat gttgagggcg atgtcgtagg
     45 aggegegeag ggeetecage accepettee gtegeggatg gegeaceaeg tacaegtagt
                                                                               180
     47 agacgtgete gttgeeetge geggteeteg geageageag eeeegtgtee geeaggeeet
                                                                               240
     49'cctcatageg gegtgecace geeeggeggg cetegatgta ggaeggeaac egggaeaget
                                                                               300
     51 tgcgccgcag gatctctgcc tgtacttcgt ccagccggct gttgtgcccg ggggtttcga
                                                                               360
                                                                               420
     53 cgaegtagta geggetetee atgeegtagt agegeageeg eegeageegg teegeeaeee
     55 getegtegte ggtgageace gegeegeegt eeeegtaege geeeageace ttggtegggt
                                                                               480
     57 agaaggagaa egeggeegeg teaceggteg ageeggegag teggeegtge eggegeee
                                                                               540
     59 cgtgcgcctg cgcgcagtcc tccaggatca ccaggttgtg ccgggcggcc agatcgcgca
                                                                               600
     61 geggtgecat gtecaegeae tgeeegtaga ggtggaeegg cageagaeae egggtgegtg
                                                                               660
     63 gegtgaggae .ggcctccacc tgggaegtgt ccatcaggta gtcctcctcg cgcaegtcca
                                                                               720
     65 cgaagacggg cgtggcaccg gccgagtcga tcgcgacgac cgtgggcgcg gcggtgttgg
                                                                               780
                                                                               840
    67 acacggtgac gacctcgtcg ccgggcccga cacccaaggc ctgtaacccc agcttgacgg
                                                                               900
     69 cgttggtccc gttgtcgacg ccgacggcat gtccgacgcc ctggaatgag gcgaactcgg
     71 actogaagec gegeaegete teaecgagga egageeggee ggageggaae acegteteea
                                                                               960
     73 eggeategtg gatgteeteg egtteeaget egtatteegg eagatagtee eacaegtgta
                                                                              1020
     75 cggtcatcga gcccctccgg gattctccct gcgaaaagtc accactctac gacaacgttc
                                                                              1080
```

77 accacteget tttteeteaa egteegettg agaeggeeeg geetgetgtg geeeggggaa

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/830,160

1140

PATENT APPLICATION: US/09/830,160

DATE: 05/17/2001 TIME: 15:34:08

Input Set : A:\Pto.amc
Output Set: N:\CRF3\05172001\1830160.raw

79 a	aggtgcggtc	gttatcatcg	actccgtctt	ctcattcgga	ggttgttcag	ggtgaaggga	1200
					cgactctcgc		1260
83 (cagettetee	ccgtcgggga	caagccgatg	atctactacc	cgctctccgt	gctgatgctg	1320
85 (gccggcgtca	cggacatect	catcatcagc	acaccgcacg	aactcccccg	aatgcgccgt	1380
87	ctattcaaca	acggcgcaca	gctcggactc	cgcctggcct	acgccgagca	ggagaaaccc	1440
89 a	aggggtatog	ccgaggcgtt	cctgatcggt	gccgaccacg	tgggaagcga	tgccgttgcg	1500
					agggggtgct		1560
					tcaaggatcc		1620
					tcgaggagaa		1680
				_	acaacgaggt		1740
					tcaccgacat		1800
					ggggattcgc		1860
					g tgtccgccct		1920
					gcatgggctt		1980
107	caggeetget	acgaactggg	cacacacata	tecaactec	gctacgggca	gtacgtgatg	2040
109	accateacaa	aggagtgcac	agagcagata	taàacaacc	g tgccgggtgg	gcgaacggcc	2100
111	concettace	caaccccaca	caccccaaca	aacaacccc	ggccggtcag	cccatcatcc	2160
					ggtactcctc		2220
					g tggcctcgcc		2280
					gggacacccg		2340
					ccaggaaccg		2400
					gegeggeeat		2460
					g agcagaacag		2520
					c cgggccgcac		2580
127	geguegaaaa	geeegeeeag	ggegaagteg	tastasass	getecaggee	atacatataa	2640
							2700
					c cgcaggcgac		2760
					a cccgttcggc		2820
				_	cegecagtte		2880
		_			acccggcgcc		2940
					a ccggcaggcg		
•					a ccacgtgcgc		3000
					ggccgaaggg		3060
					a gaagcatccg		3120
					c ccgcgtgccg		3180
					c cggtccgggc		3240
					g tgaggtccag		3300
		_			gggggacgcc		3360
					g tgccggtgtg		3420
					cgtccggaga		3480
					c cctgccgggg		3540
159	tcgtgcacca	ccgggtcggg	ctcgtactgt	. gaggaatcca	a ccggetgacga	aaggtcgatg	3600
161	agccgcacgc	gcacctccgg	gttcgtagac	: gggctcggct	t gacgcagcgc	gggtacgacg	3660
					a cgggcgggca		3720
165	cggccggcgc	cgcaccggcg	ccgtcccggc	: cgagcgggaa	a tccagggagg	gtatagcggc	3780
167	gcgccccacg	ctgccgtcat	ggtgatgaaa	ctgacggaca	a gcgagctggg	gcgtgcgctg	3840
169	ctctcgctgc	gtggttacca	gtggctccgc	ggcatccac	cacgateceta	cgccctgctg	3900
171	ctgcgcgccg	agagcgacga	teeggegeag	ctcggccgg	c tgctgcgtga	acgcggccgg	3960
173	ctccaccgca	gcgacaccgg	cacctgggtc	accgcggac	c atgcgacggc	ctcccggctg	4020
175	ctcgccgacc	cgcgcttcgt	gctgcgccgc	cegeeggeeg	g ggcccgccac	cggcaccggg	4080
	-				•		

PATENT APPLICATION: US/09/830,160

DATE: 05/17/2001 TIME: 15:34:08

Input Set : A:\Pto.amc
Output Set: N:\CRF3\05172001\I830160.raw

	5						
177	gacgtcatgc	cgtgggaaga	ggccacgctg	agcgacctgc	tgcccctcga	cgaggcgcgc	4140
179	ctgacgaccg	accgggcacg	gtgccgccgg	ctcggcgcga	ccgccgcgcg	gategeggeg	4200
				ctggccgggg			4260
183	tcaacgggcc	acttcgacct	cagggccgac	tacgccctcc	cgtacgcggt	cgagccggcc	4320
185	tgcgcgctgc	teggeetgee	ggccgggcag	tgttccctct	tcggcgcctt	ctccccggcc	4380
				cgccttccgg			4440
				cggctggccc.			4500
				ctgacggccg			4560
193	gtccacctgg	tctgcgaggc	ggtcgccgcc	ctgtcgcacg	accccgggca	ggccgggctg	4620
195	ctcagggacg	acccggtact	cgccgcaccg	gcggtcgagg	agacgctgcg	ccacgcaccg	4680
197	cccgcccgtc	tgttcaccct	ccacgcgacc	ggaccggagc	gcgtcgcgga	cgtcgacctc	4740
199	cccgcgggcg	ccgaggtcgc	cgtcgtcgtg	gcggcggcgc	accgcgatcc	ctcctggtgc	4800
201	ccggaccccg	accgcttcga	cctcaccagg	aacgagcggc	atctggcact	gccgccggat	4860
203	ctgccgctgg	gggcgctcgc	cccgctgctg	cgcgtctgcg	cgaccgcggc	cgtcgcggcc	4920
205	ctcgcggccg.	gactcctccc	gctgcgggcc	gtcggcccgc	ccgtacgacg	gctgcgtgcc	4980
				gtcgccccgt			5040
209	gtcatccccg	gcccgccttc	ccccgcccgc	aacggaaggg	actctccatg	gacaaccgcg	5100
211	agaccgtacg	accggtgagc	gtctgccggg	tctgcggcgg	caacgactgg	caggacgtcg	5160
213	tggacttcgg	tgacgttccc	ctcgccaacg	gcttcctgtc	cccggccgac	tcctacgaga	5220
				cctgccgcgc			5280
217	cccacgtggt	cgaccccgag	gtgctgtacc	gcgactacgc	ctacaccacc	cccgactccg	5340
219	aaatgatcac	ccagcacatg	cggcacatca	ccgcgctgtg	ccgcacccgt	ttcgagcttc	5400
221	ccccggacag	cctcgtcgtg	gagctgggca	gcaataccgg	ccgtcagctc	atggccttcc	5460
				accccgcgcg			5520
225	ggcgcaacgg.	catcgagacc	ttccccgact	tcttctccca	cgacgtggcc	cgcaccatcc	5580
227	ggcgcgacca	cgggcaggcg	cggctcgtgc	tgggacggca	tgtcttcgcc	cacatcgacg	5640
229	acgtgtcgga	catcgcggcc	ggcgtacgcg	aactcctgtc	tecegaeggg	gtgttcgcga	5700
				agaaggtcgc			5760
				tcgtcaccct			5820
235	gggtgctcga	cgtggagcgg	ttcggcgtgc	acggcggatc	ggtcctcgtc	ttcgtgggcc	5880
				ccgtccccga			5940
				gcacgttcgc			6000
				tcgtggccca			6060
				tcacggtgtg			6120
				agcagggcag			6180
				aacacatccc			6240
				aggagacggc			6300
		-		tcctcacgtc			6360
				gcgcagggct		•	6420
				cccgagcacc			6480
257				gcgcacccgt			6540
259				ggtggccgcc			6600
261				ggccagcgtc			6660
263				gttgacgacg			6720
265				gcggccctgc			6780
267			-	cgacagcacg			6840
269				agcgtccgac			6900
271				gtggatcagc			6960
273	geggegggee	agcagccggg	gcccgccgcc	gttgacgcgc	atcgcctccg	cccaccgcga	7020

PATENT APPLICATION: US/09/830,160

DATE: 05/17/2001 TIME: 15:34:08

Input Set : A:\Pto.amc
Output Set: N:\CRF3\05172001\I830160.raw

	· · · · · · · · · · · · · · · · · · ·						
275	ctcggcgccg	tccacgtccg	tgaaggcggc	gcagttgacc	accacccgcg	gccggtgcgc	7080
277	ggcgaacgcg	gcgtccaccg	cccgcccgtc	ggtgatgtcc	agcgcgcgcc	gcccgagtac	7140
279				ggtctccgcc			7200
281				cacgcgcatc			7260
283	acgacggtgg	acgtaccgcc	cggcgccgtg	actccccgct	tgagcggctc	ccaccaggac	7320
285	cggttctcgc	ggtaccactg	gaccgtcgag	cgcagccccg	aggagaactc	ccgcgccgga	7380
287	cggtagccca	gttcctcacg	ggccctgccc	cagtccaggc	tgtaacgcag	gtcgtgcccc	7440
289	ttgcggtcgg	gcacgtgccg	gacgctgctc	cagtccgccc	cgcacagctc	cagcaacata	7500
291	cccaccagct	cccggttgga	gagctcccgg	ccgccgccga	tgtggtacac	accgccgggc	7560
				cggcagtggt			7620
				accggcagcc			7680
297	aagcgcggga	tgagcttctc	cgggtgctga	cgcgggccgt	agttgttgga	acagcgggtc	7740
299	acccgcacgt	ccaggccgtg	cgtgcggtgg	caggcgaacg	ccatcaggtc	ggccgacgcc	7800
				agcgggtgct			7860
303	gcgatggagc	cgtagacctc	gtccgtggac	accaggacga	agggctccac	gccgtggcgc	7920
				acgacgttgg			7980
307	.gcgcggatgg	accggtcgac	gtgcgactcc	gcggcgaagt	ggacgacctg	gtcgctgtgt	8040
309	gccatcagct	cgtcgaccag	ctcggcgtcg	aggatgtcgc	cccgcacgaa	gcgcagccgg	8100
311	tcaccgcgta	ccgcgtccag	gttcgtgagg	ttgcccgcgt	acgtcagttt	gtcgaggacg	8160
				ggggcccggt			8220
315'	taggccgagc	cgatgaaacc	gaccgcgccg	gtgaccagga	tgttcacgtc	cgtcgtcgcg	8280
				ccgtcgggtg			8340
319	cctcgaccca	gcgtcggggg	cggccgtgga	ggagcggttg	agcttcggcg	cagcggcggc	8400
				ccaacggttc			8460
				agagaggag			8520
				caacaccggc			8580
		_		caccttggaa			8640
				gcgggcgacg			8700
				gtgggtcaag			8760
				gcccaccgac			8820
				gatccgcgac			8880
				gtggcccgac			8940
				tcagaccagg			9000
				tcgaaggagg			9060
				tcggagacgg			9120
				cgcgagatct	•		9180
347				gacatcctgg			9240
	_			ggactgccgc			9300
				cccgtcgaac			9360
				cggcaggagc			9420
				tacccgtggc			9480
			_	tgccgtaccg			9540
359				gtcgtctaca		-	9600
			_	cgcctgtcgc			9660
363				cggcagtgag			9720
365				ggccagcgcc			9780
367			-	ccgcacccgc		-	9840
369				ggcacccagc			9900
371	cccgatgccc	agcccggtca	gcacatcgcg	gttggtgtcc	tggtaggagc	cgtgcgggat	9960

PATENT APPLICATION: US/09/830,160

DATE: 05/17/2001 TIME: 15:34:08

Input Set : A:\Pto.amc
Output Set: N:\CRF3\05172001\I830160.raw

	373	gacgcactgc	gggacgccgg	cggccagggc	cgtcagcagt	gtgccgctgc	ccccgtgatg	10020
	375	gatgatcgcg	tcgcacgtct	ccagcagcgc	gcccagcgga	atccactcca	ccaccggtac	10080
	377	gttcgcgggc	agttcaccga	gcagggccag	gtcgccgccg	cccagggtca	gcacgaactc	10140
	379	cgcgtccacg	tccgccactt	cggagaacag	cggggccagc	ttggcgatgc	cgcccgacag	10200
	381	cgcgtcgatg	gagcccagcg	tcaccgcgat	acgccgccgg	ccggccgcgg	gcggcagcca	10260
	383	gtccggcagc	accgctccgc	cgttgtaggg	gacgtaccgc	atcggccagg	cacccgggga	10320
	385	gcgccggtcc	tccggcagca	gcgcctccac	gctcggcggt	gtcgtcgtca	gccgcacgga	10380
	387			cgccgtggcg				10440
	389	gagcgcgccg	agccccggct	cgctgtccgc	gggacccagc	ggcagctcta	cgcacggcag	10500
	391			gcgggcccgc				10560
	393	gggccgccag	ctccgcgccg	tccgcagcgc	cccgtcgacg	gccaccgccg	atacccgggc	10620
	395			cctcgcccag				10680
	397			gcttcgcgta				10740
	399	cagcccggcc	teegegaegg	cgcggatgtc	gccccccgtg	gcgtagcgga	cctcgtggcc	10800
	401	gagagcgcgc	agcgcctgtg	ccagcggcac	cgtcggcagg	atgtggctga	gcccgggtga	10860
	403	agtgatgaac	aacgcacgca	tgatgccccc	tgttcgacat	gaacctggaa	cacgcatcct	10920
	405			cggtcgacgc				10980
	407	ggggccggtc	cggccacgac	gcaggctcca	gcggacgtcg	acggcgggga	cgcagcgtgg	11040
	409	tcgccgggag	gcatcgatga	cagtattggt	aaccggagcc	acaggaaacg	tcggccggca	11100
	411	cgtcgtcacc	gggctactgg	ccgccggccg	ccgggtgcgg	gcgctgaccc	gcacacccga	11160
	413	ccggtccggc	ctgcccggcg	gcgcggagat	cacaggcggc	gacctgaccc	gcccggagac	11220
				gtgtcgaagc				11280
				cgcgacgggc				11340
				ccgacaccgg				11400
				cccatgtgcg				11460
		, ,,,,		gcgcggaggg				11520
-				ccgacgtcgc				11580
				acagcgtgac				11640
				ggctcggccg				11700
							actacgtcct	
		-		ccgagcggcc	_			11820
				cgctcgccca				11880
	437	gtgaccggag	accgcgtcca	cccccccc	acagaaaggc	gacgcccgtg	cgcgtactgc	11940
				gcccacttct				12000
				cgggtggccg		_		12060
				ccggtcggcc				12120
			-	gccctgcacg			_	12180
				ctcaaagggc				12240
				atggtcgacg				12300
				ccgttcacct				12360
				ctgtccttcc				12420
				caggagcccg				12480
				cacggccact				12540
				gcccccgtgc				12600
				ggactggtgc				12660
				ctggtcaccc				12720
				gaccttttcg				12780
				cagcgggagc				12840
	469	tcgtcgagca _.	cgtgccgctg	gacgccgttc	tgccgacctg	ctcggcgatc	greeaccacg	12900

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/830,160

DATE: 05/17/2001

TIME: 15:34:09

Input Set : A:\Pto.amc
Output Set: N:\CRF3\05172001\1830160.raw

L:14 M:270 C: Current Application Number differs, Replaced Current Application No

 $L:14\ M:271\ C:$ Current Filing Date differs, Replaced Current Filing Date

L:481 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 L:483 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1